# Attributing development impact: lessons from action research with the Qualitative Impact Protocol (QuIP)

**Vienna Evaluation Network, May 7, 2020**

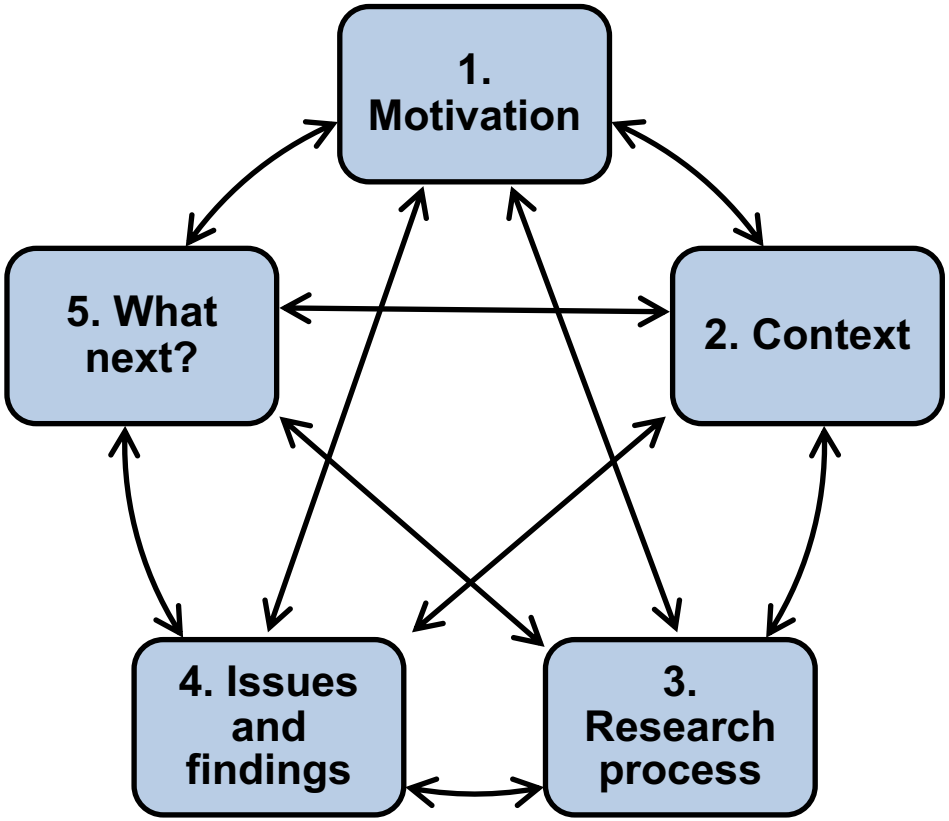
**James Copestake,** University of Bath

How do we know whether public interventions are achieving their intended purpose? Impact assessment is a necessary condition for evidence based action but evaluators struggle to do it effectively, particularly in complex and fast changing contexts. This seminar will explore this tension in relation to international development practice by reflecting on action research designing and using a qualitative impact protocol - the 'QuIP' – which aims to be flexible, credible and cost-effective. Drawing on examples of its use from Mexico to India via Mozambique and the UK the presenter will explore ways of addressing confirmation bias, cherry picking, qualitative data analysis and re-docking problems. The seminar draws on "Attributing Development Impact", available as a free e-book at bit.ly/QuIP-OA.

Bath SDR

QuIP

# Outline




Bath SDR

QuIP

# 1. Motivation – the challenge addressed

Even with effective change monitoring, how can social investors credibly assess their contribution to observed outcomes, particularly in complex contexts?

| | i.e. **with less** |
|---|---|
| Enabling intended beneficiaries to voice their views in an **ethical** way | Extractive data collection |
| Reflecting the **diversity** of their experiences | Over-simplification |
| Making **cost-effective** investments in finding out more | Obsessive measurement disorder |

Bath SDR

QuIP

# Additional research questions

How do social investors construct and adapt causal models of change in complex contexts?

What role can formal evaluation play in evidence-based *political deliberation* over these models?

How is impact evaluation delegitimized?

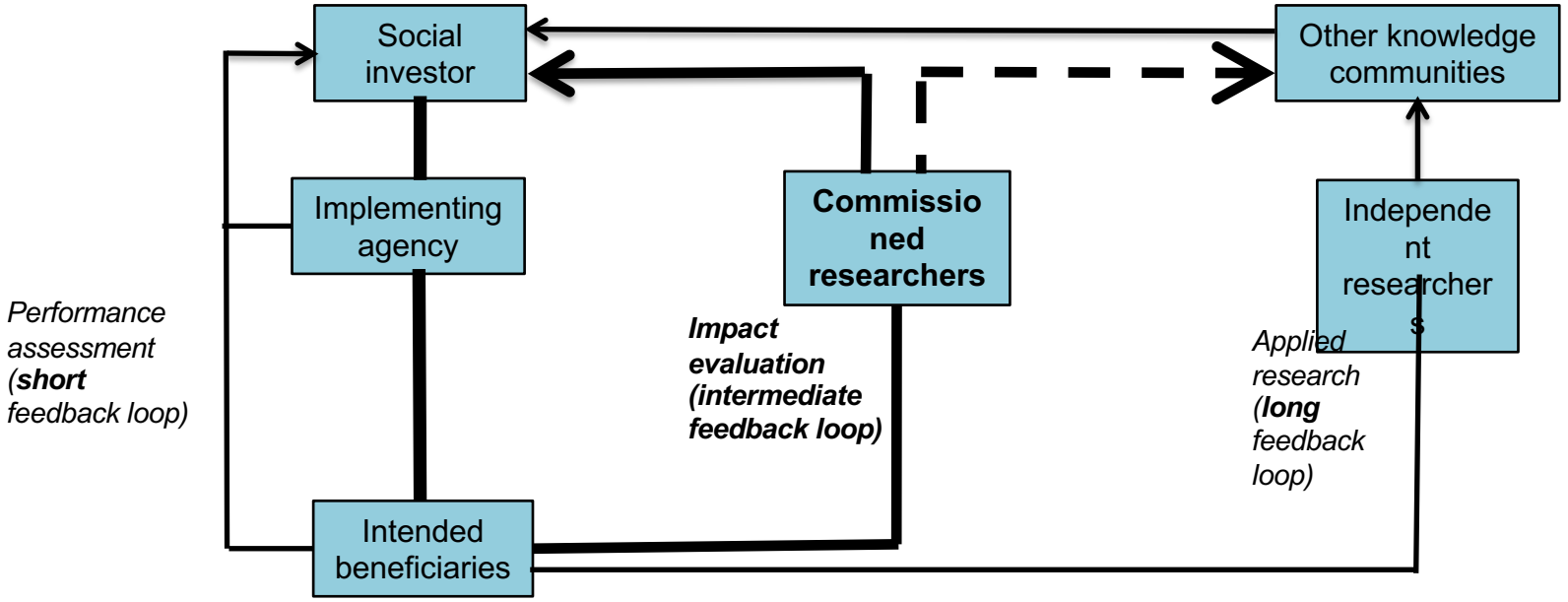
| | Validity or credibility | Usefulness |
|---|---|---|
| Change/results monitoring | Key variables can't be measured (**immeasurable**) | The cost of measuring key variables is too high (**imprudent**) |
| Theory based causal claims | Insufficient contextual data (**incomplete**) | Misdirected, untimely, insufficiently precise (**irrelevant**) |

Bath SDR

*Adapted from Molecke & Pinkse (2017)* Accountability for social impact: a bricolage perspective on impact measurement in social enterprises. *Journal of Business Venturing.*

QuIP

# 2. Context




**Project specific theories of change** – how (well) will/did this particular intervention perform?

**Mid-range theory / generalisability** - how might similar interventions perform in other contexts?

**General theory** - how does this fit in with wider understanding?

*Mission drift?*

*Obsessive measurement disorder?*

Bath SDR

QuIP

# Multiple and contested impact evaluation options

| **Rely on…** | **Issues…** |
|---|---|
| Operational data and **performance management** | Mostly change monitoring, not impact<br>Cognitive biases, vested interests & mission drift<br>Weak external credibility |
| **Variance based impact assessment** (observational or experimental, including RCTs) | Costly and indivisible<br>Can be extractive (survey slavery)<br>Narrowly framed (spurious precision, OMD)<br>What works (on average) but not how |
| **Qualitative social research** | Akerlof's lemon problem (lack of transparency)<br>Credibility and cost-effectiveness trade-offs |
| **Participatory learning and action** | Limited generalisability<br>Limited credibility to non-participants |

Bath SDR

QuIP

# What a QuIP does not provide

| *Does not provide…* | *Responses* |
|---|---|
| Estimates of the magnitude of average treatment effects | Use as one input into microsimulation<br>Run alongside a quantitative impact evaluation |
| Statistically representative frequency counts | Reveals scope and range of responses<br>Combine with Bayesian updating<br>Use alongside quantitative surveys. |
| Objective 'facts' | Perceptions matter!<br>Incorporate with other approaches |
| Recommendations for action | Combine with process evaluation<br>Address in follow-up stakeholder engagement. |

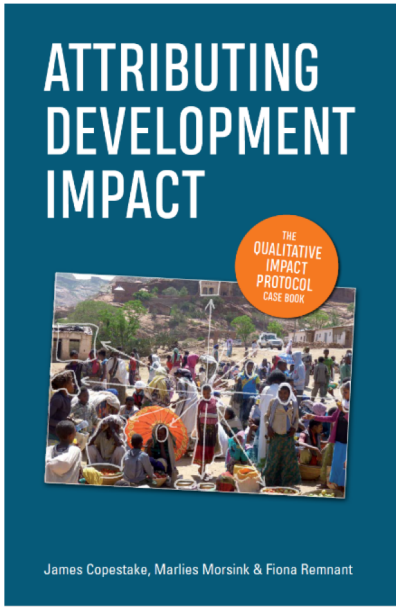
Bath SDR

QuIP

# 3. The research process

**Design and pilot testing** (2012-15)

Grant funded collaborative action research to design and test a qualitative approach to impact assessment – tested in Malawi and Ethiopia.




**Commercial testing** (2016-)

Set up Bath SDR Ltd as a social enterprise to deliver QuIPs in a wider range of contexts. 40+ commissioned QuIP evaluations in 20 countries.

**Continued action research** (2016-)

Ongoing learning and publication on how to do impact evaluation better, including "*Attributing Development Impact: The QuIP case book*" (2019) - **bit.ly/QuIP-OA**

Bath SDR

QuIP

# BSDR QuIP studies 2016-2019

## Activities

Child nutrition
Climate change adaptation
Community mobilisation
Early famine response
Factory working conditions
Housing improvement
Medical & midwife training
Microfinance
Rural livelihoods
Value chain improvement
Sexual & reproductive health rights
Organisational development

## Countries

Bolivia
Burkina Faso
Cameroon
Ethiopia
Ghana
Kenya
India
Indonesia
Malawi
Mexico
Mozambique
Nepal
Pakistan
Sierra Leone
Tanzania
Tajikistan
Uganda
UK
Zambia

Bath SDR

## Commissioners

Acumen
Bristol City Council
C&A Foundation
Diageo
Self Help Africa
Habitat for Humanity
Oxfam
Save the Children
Seed Global Health
Tearfund
Tree Aid
Rutgers Int'l
Itad
Concern Worldwide
Aga Khan Foundation
MannionDaniels
Send a Cow
Oxford Policy Mgt
Power to Change
Opportunity Int'l
AgDevCo

QuIP

# An overview of the QuIP

If you haven't done so already then please click

here

and listen to the short video

# Features of the QuIP

**Self-reported attribution** (with latent counterfactuals) Not statistically inferred attribution based on exposure variation.

**Eclectic** (contribution analysis, goal-free evaluation, outcome harvesting, most significant change, process tracing, realist evaluation…)

**Dual purpose**: Exploratory (theory building) and Confirmatory (theory testing).

**Bayesian in spirit** – augmenting prior theory incrementally; never starting with zero understanding.

Bath SDR

Click here

QuIP

# More features of QuIP

Designed for **complex contexts** (including presence of multiple causal drivers, measurement problems, rapid change and possible unintended consequences).

Focuses on making **qualitative data analysis** structured, transparent, accessible and useable

Uses **data visualisation – causal maps -** to make reported outcomes more digestible

**Good enough** – balancing credibility and cost-effectiveness; not aiming for absolute or universal truth.

Bath SDR

QuIP

# Combining QuIP with other methods

**Supplementing** operational data and observation, and process monitoring.

**Substituting** for other impact evaluation approaches.

**Integrating** QuIP with other methods (see below)

**Before** - For exploratory analysis, scoping and identifying variables for quantitative studies.

**After** - To confirm of deepen understanding of impact claims based on quantitative studies

**Alongside** – To identify mechanisms to explain correlations, instead of relying on theory or speculation.

As an **input** into quantitative analysis (e.g. to inform modelling and simulation).

Bath SDR

QuIP

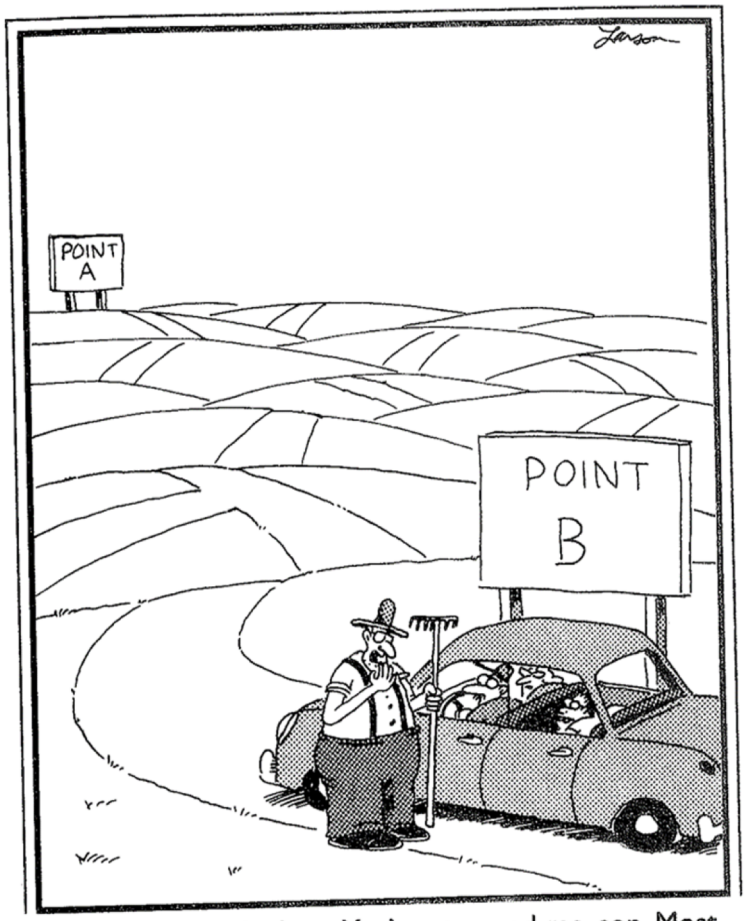
# 4. Key methodological issues

| Issue | Problem | Solutions |
|---|---|---|
| Self-reported attribution | **Confirmation bias** | Double blindfolding |
| From field data to synthesis | **Opaque data analysis** | Thematic (inductive) and attribution (deductive) coding.<br>Dashboard visualisation and cross-referencing numerical & text data. |
| Robust generalization | **Cherry picking** of cases and sources | Transparent case selection informed by large 'n' data on context, outcomes (ideally) and prior theory of change. |
| Effective data use | The **re-docking problem** | Clarity with the commissioner at the planning stage over scope for deliberation |

Bath SDR

QuIP

# Blindfolding




"Well, lemme think. … You've stumped me, son. Most folks only wanna know how to go the other way."

**Why do it?**

To reduce pro-project, framing and confirmation biases

To give equal weight to all possible drivers of change

**Is it ethical?**

Informed consent

Time-bound

Greater good (should be proportionate)

**How far to go?**

A design choice – never zero nor 100%

Only one way of reducing bias

Can be temporary

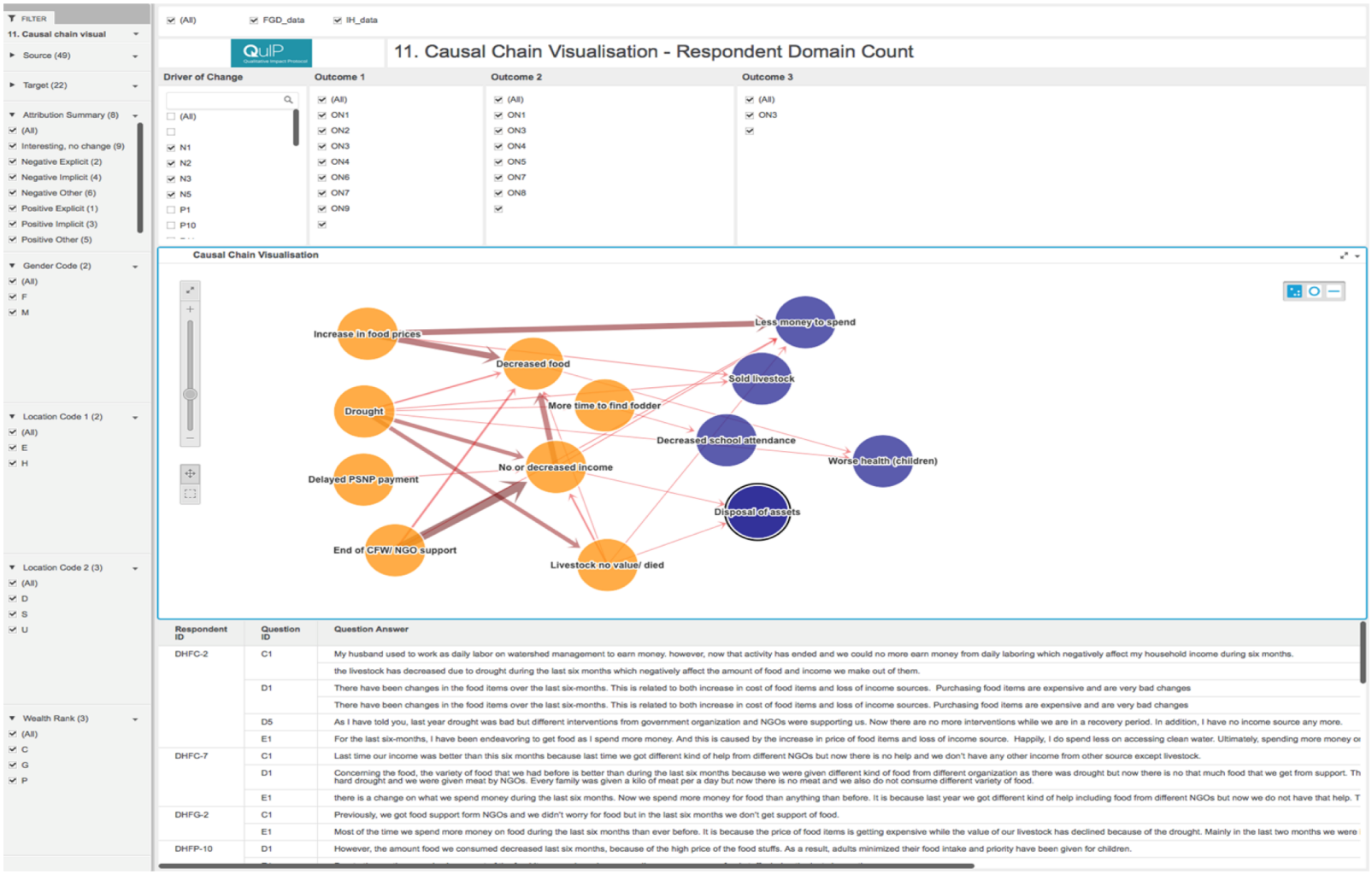
Can be combined with other sources

Exploratory analysis can also be blindfolded.

Bath SDR

QuIP

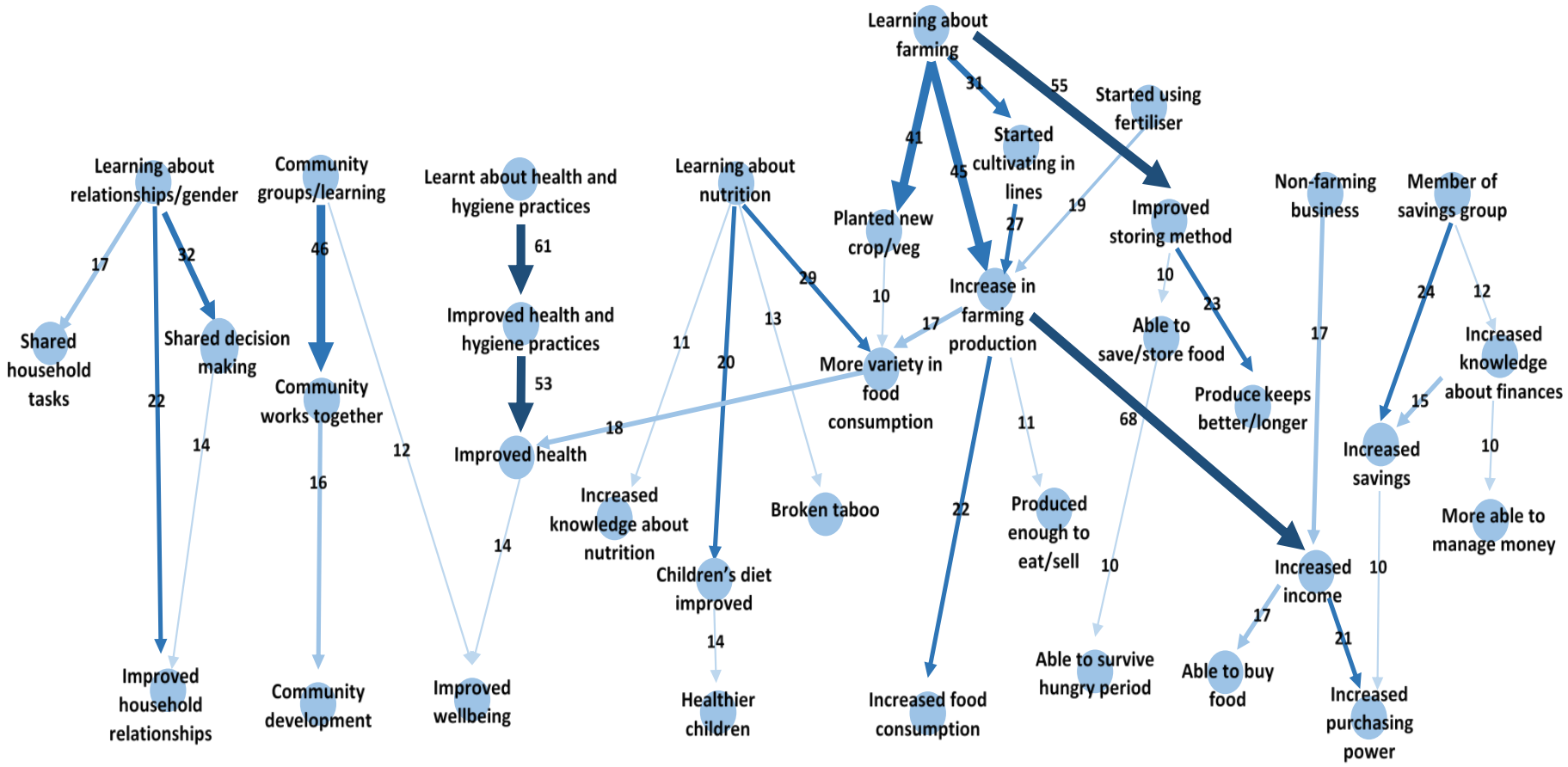
# Thematic coding and visualization




Bath SDR

QuIP

*Citation Count: 10+*




Bath SDR

Extract from a QuIP study for Save the Children of an integrated rural development programme in Mozambique

QuIP

# Evidence selection – definitions and principles

**Saturation –** Maximize potential to learn about the fullest possible range of drivers of change affecting the selected population (*Exploratory)*

**Bayesian updating -** Maximize the potential to test prior theory (*Confirmatory*)

**Heterogeneity and differential impact -** Capture the most important sources of variation (see below).

**Equating marginal benefits and costs –** Invest in more data until the extra evidence does not justify the extra cost.

- The principles for doing this are NOT the same as those for estimating the average value of a known variable across a population
- Random selection is NOT necessarily best, indeed is unlikely to be.
- Doing all data at the same time is NOT necessarily optimal.
- There is no scientifically optimal sample size, but transparent and reasoned case selection is important to rebuff criticisms of 'cherry-picking'.

Bath SDR

QuIP

# Evidence selection options (Linked to availability of prior data)

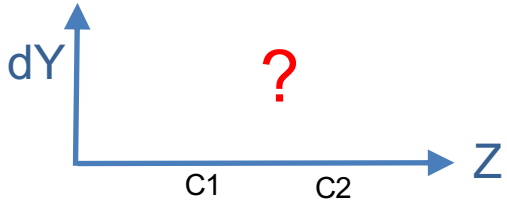
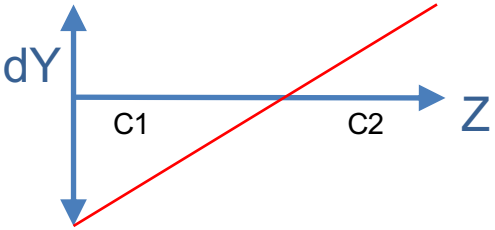
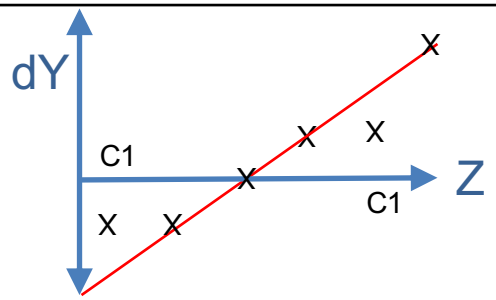
| Option | Treatment data (X)? | Outcome data (Y)? | Contextual data (Z)? | Comment |
|---|---|---|---|---|
| A | No | No | No | Random selection across full population is the only option |
| B | Yes | No | No | Select randomly from quota samples across categories of treatment or exposure |
| C | No | Yes | No | Select purposively to include positive and negative deviants |
| D1 | No | No | Yes | Select purposively to reflect important dimensions of variation across the population (e.g. gender, age) |
| D2 | No | No | Yes | Select purposively to include likely positive and negative deviants according to prior theory. |
| E | No | Yes | Yes | Select purposively to include anomalous cases poorly explained by prior theory linking Z and Y. |

Bath SDR

QuIP

# Case selection – illustration

(Two cluster source data selection scenarios for a uniform X)

| | | | |
|---|---|---|---|
| 1 |  | No data for dY (outcome) or theory of change linking it to context (Z) | Maximise dispersion of cases across Z, or select randomly |
| 2 |  | Theory of change linking Z to dY, but no data for dY | Choose expected negative (C1) and positive (C2) **deviant** cases |
| 3 |  | Data for Z and dY permits empirical testing of the theory | Select on **anomalous** cases or outliers relative to theory: positive (C1) and negative (C2) |

Bath SDR

QuIP

# The redocking problem

## Work back from use

**To stimulate internal learning** (and consultation with a wider diversity of stakeholders)

**To influence operational decisions** (e.g. mid-project reviews and adjustment, scaling up, closing or adapting activities)

**To support wider deliberation** over core theories of change

**For external reporting** (with need for sufficient credibility to match potential push back)

## via outputs

A written **report** of findings

An interactive **dashboard**

**Integration** with other sources of evidence

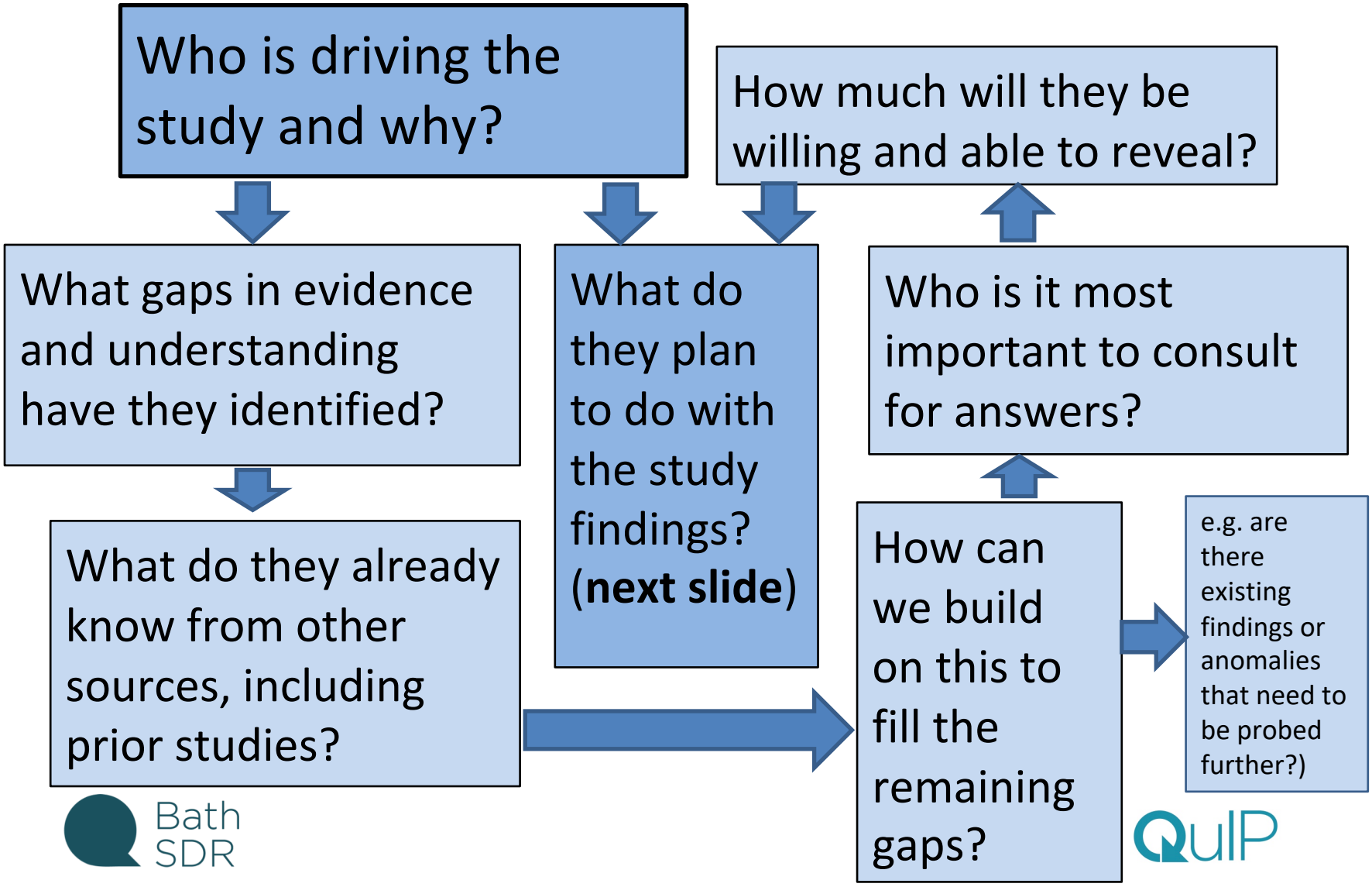
Participatory sense-making **events**

**Media** based wider dissemination

Bath SDR

QuIP

# The key to re-docking – initial framing



# What next? Evaluating complex problems

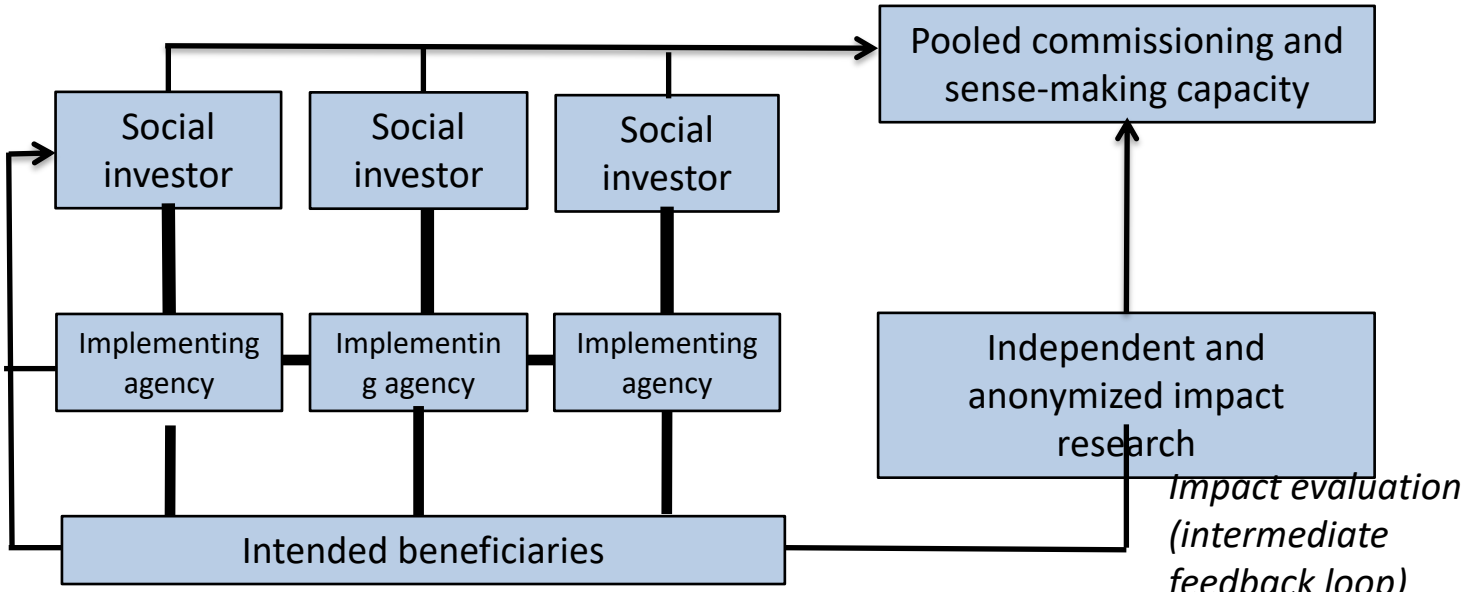
1. Elaborate the intervention **theory of change**
2. Monitor changes in key outcome indicators and their possible causal drivers across the population (**large 'n'** or wide-but-shallow data).
3. Ask key informants to identify credible causal mechanisms between observed changes (**small 'n'** or narrow-but-deep data).
4. Collect addition specific evidence to adjust these prior causal claims.
5. Iterate between these steps until a threshold of sufficiently credible evidence is achieved (or we run out of time and money).

Bath SDR

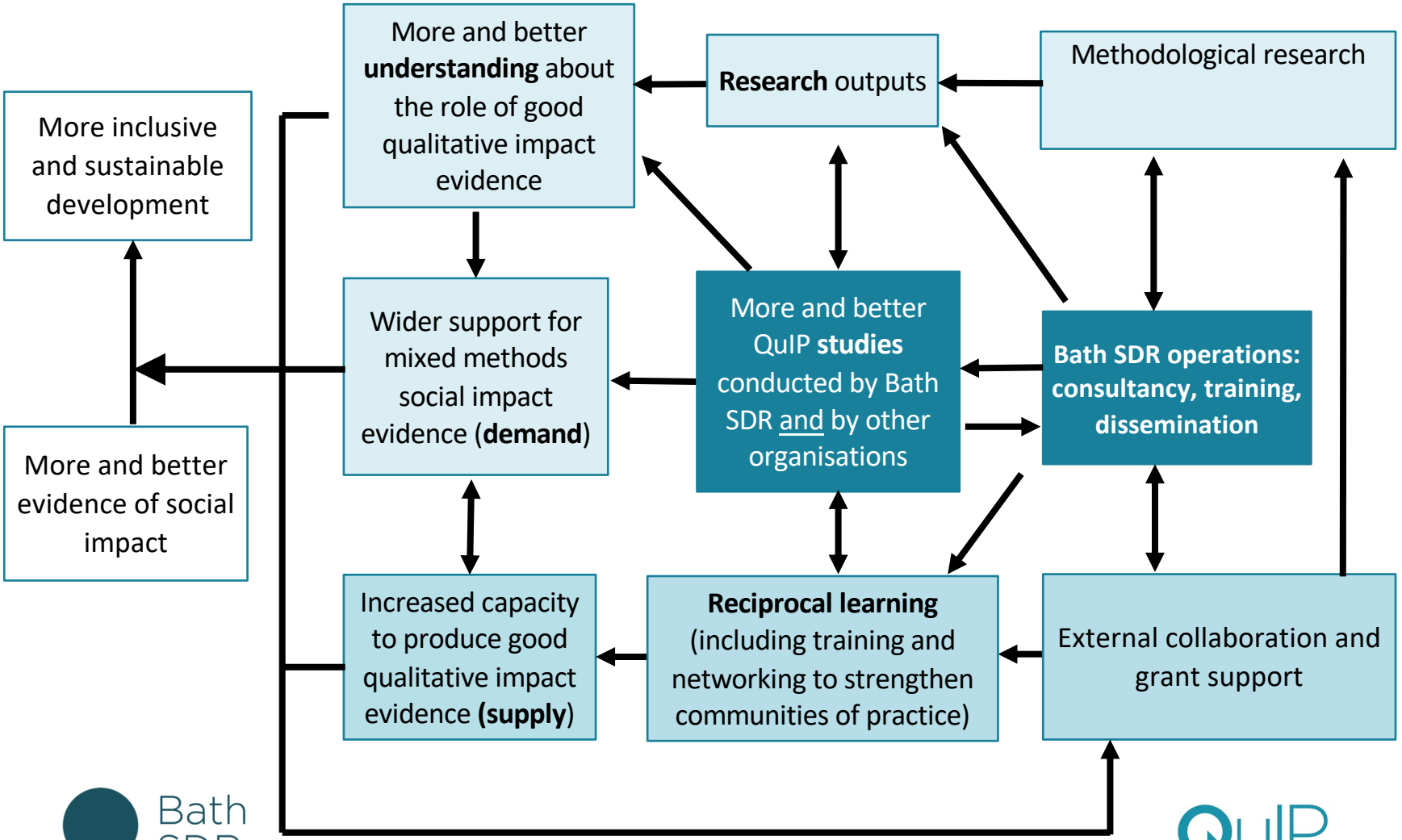
e.g. The example of fire detection

QuIP

# Pooled commissioning to address the silo problem



# Bath SDR: QuIP theory of change




Bath SDR

QuIP